

PRINCIPLES OF ASSESSMENT

Mrs. JYOTI S. MOHITE
Associate Professor,
Dept. of Audiology and Speech Therapy,
T. N. Medical College, B. Y. L. Nair Hospital.
Email: jmohite70@yahoo.com



Students Assessment



METT Cell

T.N.Medical college & B.Y.L.Nair ch Hospital

- Whether practicing in a rural, community, or an academic setting, physicians from all clinical specialties will participate in assessment.
- These assessments may be for trainees, peers, and more recently, for self-assessment.

At the end of the session, the participants should be able to

- Review purposes of assessment
- Discuss characteristics of an effective assessment
- Distinguish between summative and formative assessment
- Enumerate different methods of assessment
- List factors governing choice of assessment methods

What is the need of Assessment????

Discuss in groups
Need for **students & teachers**

Need for Students



- Gives idea of his standard of learning
- Can take necessary corrective steps to improve
- Gives idea of where he stands in the class
- Whether able to deliver during exams
- External motivator

Need for Teachers



- To assess the ongoing learning program
- To modify the ongoing program
- To identify common misconceptions
- To calculate internal assessment

What is the need of Assessment????

General

- To certify – pass/fail
- To promote
- To select
- To rank
- Feedback

Types of Assessment

- Formative

Done on a regular basis during the course

Eg Periodicals, Terminals

- Summative

Done at the end of the course

Eg university exams

Criterion	Formative	Summative
Purpose	Detect specific strengths / weakness of students	Determine overall Achievement
Frequency	Ongoing	At the end
Area covered	Single unit	Course content
Feedback to students	Immediate	Pass or fail
Feedback to faculty	Identify areas of weakness in instruction	Overall view of pass/failure

jsm/mett/13aug2014

Methods of Assessment

- Written – Multiple Choice Question (MCQ)
Short Answer Question (SAQ)
Brief Answer Question (BAQ)
Long Answer Question (LAQ)
Modified Essay question (MEQ)
- Viva
- Practical – Tests
Cases
- Newer - OSCE & OSPE
Portfolios
Multisource 360⁰
Mini Clinical Evaluation Exercise - Mini CEX

- *“Assessment in Medical Education addresses complex competencies and thus requires quantitative and qualitative information from different sources as well as professional judgment.”*

---- Ann Burke

- In a simplistic sense, the purpose of assessment is to enhance learning
- ‘Assessment’ covers a range of testing methods, including examinations.
- Assessments can be local or national.

Principles of Assessment

- The purpose of each assessment procedure should be clear, well understood and well communicated.
- There should be a consistent and effective relationship between the assessment process and the outcomes that need to be achieved.

- The format and design of the overall assessment system and its methods should be appropriate to what is being tested, such as:
 - a clinical skills;
 - b knowledge and decision-making;
 - c interpersonal (communication) skills; and
 - d competence in particular areas.

- Domain to be tested

Cognitive

Psychomotor

Affective

Written

Practical

Observation

Viva

Checklist

- Educational objectives
- Feasibility
- Level and ability of learners

The specific assessment methods should:

- have appropriate content and methods;
- be valid;
- be reliable;
- be based on evidence; and
- be assessed against best practice in other settings and other countries.

- The assessment should be fair, clear, accurate and well validated (based on evidence about what works).

- People carrying out assessments should be fully trained and competent, and should receive regular and frequent feedback on their performance.
- While it is rare for serious mistakes or misjudgements to happen during assessments, the process should allow for students and trainees to give feedback and appeal.
-

- If the purpose of assessment is to enhance learning, the purpose of teaching is to facilitate it.
- Before any particular teaching method can be widely implemented in health sciences education, however, there must be a method to assess its product.

Content of assessment

- Broadly classified, educational objectives fall into three domains: knowledge, skills, and attitudes.

According to Harden,

- **Knowledge objectives** - addresses cognitive domain . These range on a continuum from being able to recall factual events to integrating processes for problem solving.
- **Skills objectives** - involve psychomotor aspects that are needed to be an efficient clinician.
- **Attitude objectives** relate to personal qualities of the learner and their approach to medicine, patients and their peers.

Attributes of an effective assessment?

- Validity
- Reliability
- Accountability
- Flexibility
- Comprehensiveness
- Feasibility
- Timeliness
- Relevance to both the examiner and examinee.

Validity

- Validity – “it measures what it purports to measure”.
- face, content, construct, and criterion validity.

- **Face Validity**
- Face validity, the most subjective form of validity
- Expert opinions

- **Content Validity**
- Content validity measures the extent to which assessment items reflect the overall domain of knowledge and skills required for mastery of the subject.
- Content validity is also highly subjective in that it relies on the opinions of content experts about the relevance of items used.

Construct Validity

Construct validity examines the extent to which an assessment measures a non-observable trait that explains behaviour.

- **This allows an assessor to infer a psychological construct from test scores.**

- **Eg: one may theorize that resident performance in the intensive care setting relates to a sophisticated understanding of animal physiology.**
- **difficult to assess every aspect of animal physiology (i.e. from single celled organisms to humans),**
- **a validated assessment in general animal physiology could be administered and correlated with established intensive care test scores.**
- **If correlation is high, the relationship is demonstrated.**

- **Criterion Validity**

- Criterion validity examines the degree to which a tests correlates to other measures of performance.
- Two subtypes: Concurrent and predictive validity.
- Concurrent validity is the degree of agreement between the scores of the index test and an established one (i.e., the ““Gold Standard””) when administered at the same time.
- Eg: to improve time and cost efficiency by introducing computer simulators to test gross anatomy knowledge.

To assess concurrent validity,

- administer both the computer simulation and the traditional cadaver examination.
- If correlation with the examination scores is high, the simulator demonstrates the validity of the established assessment.
- With this agreement, one can comfortably adopt the new assessment and utilise its advantages.

- Predictive validity.
- is the most likely to provide a clinically useful assessment.
- It is defined by the extent to which **test scores** relate to **actual ability**.
- Eg: test with high predictive ability would be a surgical skills simulator assessment whose scores correlate directly to performance in the operating room.

Reliability

- Reliability relates to the precision, stability or reproducibility of an assessment tools results.
- Simply stated, reliability is a term that covers the **dependability** of an assessment and measures the extent to which a test will yield the **same result** after **multiple administrations** under the **same conditions**.

- Like validity, there are a number of methods to establish a test's reliability.

Important methods include

- internal consistency
- test-retest
- equivalent forms
- inter-rater reliability.

- **Accountability**
- Any assessment mechanism must be accountable to all 'stakeholders' involved.
- To facilitate accountability an assessment must be defensible and able to provide a logical analysis or explanation of results.

- **Flexibility**

- Clinical medicine is practiced in a diverse and sometimes unpredictable environment.

Therefore, the chosen assessment method must be flexible and allow the examiner to evaluate the complete clinical spectrum of the content domain in question multiple times and in multiple settings

- (e.g., elective and emergency surgeries).

- **Comprehensiveness**
- To be effective overall, an assessment will evaluate **all pertinent objectives** and document corresponding examinee performance for the course it was designed to evaluate.

- **Feasibility**
- To facilitate acceptance by all stakeholders, an assessment should be portable, cost-effective, practical, and limit physical and human demands.
- Examinations such as the Objective Structured Clinical Examination or Objective Structured Assessment of Technical Skills (discussed later) are examples of resource intensive but valuable assessment tools.

- **Timeliness**
- **To maximise its function, assessment should be administered as close to the target behavior as possible.**
-
- **Undue delay allows for recall of target events to degrade and thus increases the subjectivity of assessment.**
- **Also, the results of the assessment should be communicated to the examinee (and other relevant parties) promptly.**
- **Failure to do so deprives stakeholders of the assessments full utility (e.g., feedback or curriculum planning functions).**
- **Indeed, if documentation is delayed, assessment is less effective as a learning tool, more subject to bias, and less defensible.**

- **Relevance**

- To be effective, the importance of the assessment must be apparent to all involved stakeholders.
- The results of assessment, favorable or not, must be used to facilitate learning and influence promotion and curriculum planning decisions.
- An assessment that is viewed as irrelevant cannot fulfill these functions because its results appear meaningless and unusable.
- Eg : LAQ on Small pox.

Case Example- resident ratings

- Her internal examination score is average,
- she scored above her peers for the Objective Structured Clinical Examination (OSCE),
- NICU attending was rather negative,
- and the rest of her evaluations from faculty are average.
- However ward posting assessment is below average.

- How are you to know how to weigh all of this information and make any kind of reasonable, informed decision about your intern?
- Most of the faculty - quiet, performs slightly above average, and should be promoted.
- One faculty member, however, says he is concerned about the intern's ability to synthesize information to formulate a reasonable assessment.

- How do I know which of these assessments of my resident are valid?
- What does validity mean?

- There are five sources of evidence that can be used to support validity:
- Content
- Response process
- Internal structure
- Relationship to other variables
- Consequences

- **Content: Does the assessment cover the areas it intends to? Factors that influence this are:**

adequate and representative sampling of the knowledge and/or skills being observed and judged.

The resident

- in the case may be receiving discrepant assessments because one or more of the assessments may not be assessing what is supposed to be assessed.
- For example, the NICU attending may not like residents to be quiet and shy, and may be assessing demeanor rather than patient care abilities.

Response Process: This term refers to the evidence of data integrity and quality control of score reporting and accuracy.

- Is the instrument's result accurately reported?
- Is the form of assessment familiar to the learner?
- Is there a system that assures the correct resident name is on the assessment tool, or are faculty mixing residents up?
- Did the resident know she was being assessed ?
- Do residents understand how to approach an OSCE encounter?

- **Internal Structure:** This refers to reliability, and issues with rater variability fall into this category.
- **raters had no guidance** on how to use the assessment instrument?
- raters are **not standardized**, then the results of the observations may not be measuring what should be assessed.
- **uniform rating tool, and the training of raters through applying the tool to a common observation, such as a videotaped encounter.**

- **Relationship to other Variables:**
- **How well does the assessment tool correlate with the outcome of another**
- **measure of the same skill?**
- **For example, does the resident who scores highly on an OSCE station that is purported to assess communication skills also receive above average assessments in clinically observed interactions?**

- **Consequences: This considers the impact of scores on learners**—whether or not the assessment is highstakes or low-stakes.
- As an example, did the resident think that if she did poorly in the NICU rotation she would fail the whole year, and thus suffer from performance anxiety during observations?
- Conversely, did she think that the NICU rotation was unimportant?

Conclusion

- Assessment in medical education is a multi-faceted and dynamic process.
- While outwardly complex, sound appraisal is centered on basic principles that allow accurate, efficient and meaningful determinations of mastery.

- As medical teachers, we strive to assess our trainees in a thoughtful and fair manner, both for their benefit and that of the patients who rely on their competence

- To be able to **select appropriate assessment tools** that measure aspects of performance that are meaningful, one must have an **informed understanding of strengths and weaknesses** of various assessment tools and techniques.

- Key concepts that need to be considered when assessing competence are *validity, reliability, and the utility or usefulness* of a given tool when it is used for purposes of assessment

References

- A Vergis, K Hardy. *Principles of Assessment: A Primer for Medical Educators in the Clinical Years*. The Internet Journal of Medical Education. 2009 Volume 1 Number 1.
- Schwartz A, editor. 2011. *Assessment in Graduate Medical Education: A Primer for Pediatric Program Directors*. Chapel Hill, NC: American Board of Pediatrics.

Thank-you

jsm/mett/13aug2014

jsm/mett/13aug2014

jsm/mett/13aug2014

Selection of assessment technique

- When choosing a method, it is important that the assessment technique be closely related to what one is trying to examine.
- For example,
- if the aim is to examine a candidate's factual recall (““knows””), a multiple choice or extended matching item examination may be sufficient.
- If a candidate's thought process is the target (““knows how””), an essay format or oral examination may be useful by allowing a free and extended arena to formulate a response.